DEPARTMENT OF DEFENCE
DEFENCE SCIENCE & TECHNOLOGY ORGANISATION
DSTO

IP Convergence in Global Telecommunications

-

Voice Over Internet Protocol (VoIP)

Ian Zahorujko, Alfred Reynolds and Bill Blair

DSTO-TR-1039

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

IP Convergence in Global Telecommunications

-

Voice over Internet Protocol (VoIP)

Ian Zahorujko, Alfred Reynolds and Bill Blair

Communications Division
Electronics and Surveillance Research Laboratory

DSTO-TR-1039

ABSTRACT

A key application on any converged global network will be voice telephony. On an Internet Protocol (IP) network, this is given the generic title Voice over IP (VoIP). This paper examines the motivation behind VoIP and the standards being deployed in support of the application. It discusses the factors that determine the voice quality to users, and measures that can be made. The impact of VoIP on Speakeasy is given some limited consideration.

RELEASE LIMITATION

Approved for public release

DEPARTMENT OF DEFENCE
DEFENCE SCIENCE & TECHNOLOGY ORGANISATION
DSTO

20010801 081

AQ F01-10-2169

Published by

DSTO Electronics and Surveillance Research Laboratory
PO Box 1500
Salisbury South Australia 5108 Australia

Telephone: (08) 8259 5555
Fax: (08) 8259 6567
© Commonwealth of Australia 2000
AR-011-569
September 2000

APPROVED FOR PUBLIC RELEASE

IP Convergence in Global Telecommunications

-

Voice over Internet Protocol (VoIP)

Executive Summary

The DSTO report series "IP Convergence in Global Telecommunications" sought to illuminate developments in the carrier environment relating to the widescale fielding of Internet Protocol (IP) networks - the Internet. This particular paper examines a high profile application that has been a significant driver as well as a significant consequence of this adoption of IP. From the perspective of the user, IP Telephony, or Voice over IP (VoIP) is seen as a major theme of the rollout of IP infrastructure. VoIP will make significant demands upon the network performance.

The paper examines the motivation behind VoIP and the standards being deployed in support of the application. It discusses the factors that determine the voice quality to users, and measures that can be made. The voice quality is primarily determined by the effects of network performance, thus the VoIP application makes significant demands on the network to provide requisite quality of service (QoS). Another report in the series addresses the techniques available to address QoS.

A key difference between Defence users and the general public is the requirement for security. While IP security mechanisms is the topic of another report in this series, this report gives some limited consideration to the impact of VoIP on the operation of the current voice security device - Speakeasy.

The report attempts to educate the capability development, acquisition and operational management of Defence networks on VoIP. The report does not seek to consider whether the VoIP technology is relevant or appropriate to Defence networks.

Authors

Ian Zahorujko
Communications Division

Ian Zahorujko is a Senior Professional Officer (class C) in Information Networks Group of The Defence Science and Technology Organizations (DSTO) Communications Division, having joined DSTO 18 years ago. He has been involved in Defence Communications in the areas of RF engineering, signal processing and signal analysis. Recently he became involved in Network Architectures to broaden his experience in the field of Secure Communications.

Alfred Reynolds
Communications Division

Alfred Reynolds is a Research Engineer in Network Architecture Group of the Defence Science and Technology Organisation's (DSTO) Communications Division having joined DSTO after completing a Systems Engineering Degree at the Australian National University. He is involved in research into tactical communication systems.

Bill Blair
Communications Division

Bill Blair is a Senior Research Scientist in the Network Architecture Group of the Defence Science and Technology Organisation's (DSTO) Communications Division having joined DSTO from the Australian Army. He is involved in research into networking techniques to control Quality of Service (especially Asynchronous Transfer Mode - ATM) for military strategic and tactical communications.

Contents

1 Introduction

Voice over Internet Protocol (VoIP) is an innovative new application emerging on the Internet. This report seeks to examine the motivation behind these developments and discuss the standards that support VoIP (and indeed video and fax over the same standards).

VoIP is significant, as it is arguably the first application over Internet Protocol (IP) that is truly real-time and requires the network to meet demanding Quality of Service (QoS) performance. We already have Internet radio and TV station transmissions, which can be heard and viewed using free software, like RealPlayer or MediaPlayer on our PCs. The difference here is that while these applications require a continuous stream of data, there is no need for real-time interactivity. The underlying characteristics of the IP network were not designed to provide real-time data flows. While interpretation of the term 'real-time' can vary considerably, for VoIP it translates to reproducing the QoS that a normal telephone call offers today. Simply put, this implies virtually no clipping of sounds, high intelligibility and no perceptible delay of a user's speech in reaching the listener. There are other issues, such as security, billing, scalability and robustness, which the telephony industry must integrate with existing systems, in order to support widespread use of VoIP.

There have been recent arguments to discern a difference between "VoIP" and "IP Telephony". It is argued that VoIP is the technology that allows voice calls over an IP infrastructure whereas IP Telephony is a more holistic term covering value-added services. This paper will however use the term VoIP as the generic title for the entire topic.

VoIP offers a way for carriers to converge their voice and data networks. It also will let them value-add voice services to add functions such as web based call centres. For the public and enterprises, VoIP offers a way to reduce telecommunication costs by combining telephone and computer networking as well as integration of equipment into single units.

This report will cover the uses for VoIP, the various VoIP technologies and the issues that must be addressed for VoIP to work.

2 Computer Telephony Integration

Computer Telephony Integration (CTI) is a concept that predated VoIP. It is essentially the melding of the PC and voice services. Before VoIP, this involved creating unique ways to interface PCs to the Public Switched Telephone Network (PSTN) and a means of constructing PABX-like switching capabilities and services using the standardised, open PC environment. The emergence of "packet voice" standards, especially VoIP, has offered the CTI developer another approach to combining the PC and voice services.

The main use for CTI is in providing value-added services to customers, and automating tasks. All of these applications augment the existing PSTN network rather than creating a new system. Items that CTI implements are:

- Voice mail
- Digital Dictation
- Automated Attendant
- Interactive Fax
- Pay-per-call
- Inbound Call Centres
- Outbound Call Centres
- Transaction Processing

3 VoIP Scenarios

There are many different aspects to VoIP, and the functions considered foremost depend on what the goals of the implementor are. One way to bring order to the features is to map each solution along two dimensions:

- The core value of the solution, going from simple dial up through value-added applications
- The venue for implementation (for example, the carrier or the enterprise)

The figure below illustrates this categorisation graphically, with one of the dimensions on each axis, and representative applications in each of the four quadrants. Note that the dimensions are continuous ranges, not discrete values.

Table 1 VoIP Solutions Classified by Implementor and Value-Adding traits

Services	Internet Call Waiting Messaging	Voice-enabled Web Pages Teleconferencing
Basic Dial Tone	Real Time Fax Alternative Long-Distance	Toll Bypass
	Carrier Provided Provided	Enterprise

3.1 VoIP for Carrier Customers

VoIP promises a revolution in the private home. The convergence of data and voice services will enable home users to have one line into their homes, and they will have access to a multitude of services simultaneously.

When a normal home user is connected to the Internet via a computer modem, their line cannot accept incoming calls. Telstra and others are currently implementing a service called "virtual second line"[1] (VSL), which will let home users receive telephone calls via VoIP whilst maintaining their connection to the Internet. The system uses a program running on their PCs to emulate the telephone or to interface between the normal telephone handset and the PC. Calls to a home user, whose home line is busy because they are connected to the Internet, are automatically redirected by the carrier to a VoIP connection. Similarly, the home user can initiate a call into the normal PSTN using VoIP in the first hop to the telephone exchange.

A significant application in terms of the home user is the web based call centre. Current E-commerce systems are very impersonal and difficult to use for the new user. Currently, companies have separate phone numbers for customer assistance and a home user (in the absence of VSL) is required to disconnect from the Internet and phone this number to get help. With VoIP integration into the browser, it is possible for a user to click on a web page link to talk with an assistant. Due to the data channel capabilities of VoIP it is possible to present the operator with the calling customer's details before the call is even picked up. It is also possible for the customer assistance operator to collaborate with the home user in navigating the web site. This enables the operator to lessen the time that the call takes, making the operators more efficient and enhancing the customers' experience due to the quicker response times.

3.2 VoIP in the Enterprise

3.2.1 Motivations

Dataquest has predicted that less than 10 percent of all enterprise voice networks will be packet-based by 2002 [2]. Despite this low adoption rate VoIP offers many benefits to the enterprise. VoIP would be implemented in the enterprise for two fundamental reasons, economic advantages and access to better functionality. Some specific examples are:

- Where the cost of aggregated services (by placing voice traffic over data services acquired from carriers) is less than the sum of individual costs.
- Reduced need for PABX facilities.
- Branch office/home office voice network integration.
- Web based call centres.

3.2.2 Technologies

VoIP in the enterprise is mainly limited to internal use because of the lack of QoS in the Wide Area Network (WAN) and Internet environment. QoS is acceptable in the LAN because bandwidth is more plentiful and traffic loading is more predictable. VoIP is used as a replacement to the standard office private exchange (PABX) system. There are many reasons for doing this, ranging from only wanting to lay one cable around the office to wishing to use the extra capabilities such as video conferencing. Similar to a PABX, VoIP also gives the end user a richer set of value added services as compared to the current PSTN network. End user features from VoIP include call forwarding, call conferencing, "follow me" telephone numbers and voice mail. Moreover, with the modular construction of VoIP services, as new applications are invented it will be straightforward to implement them on the current technology.

The current model for VoIP in the enterprise is that of islands of VoIP networks operating on the corporate Local Area Network (LAN). All geographical locations of the enterprise are then connected together with voice being transmitted over the corporate IP WAN. For these calls to leave the enterprises network however they must somehow interface onto the PSTN network. This is achieved by using gateways at one or more LANs.

3.2.3 Interfaces

There are two approaches for end users to interface into a VoIP network.

- The first is via a multimedia PC using client software. Client software is a special multimedia program that operates on the users PC. The client uses the sound card to read and write audio data, and it then encodes this into a VoIP format so the data network can transport it.
- The second method is via VoIP handsets. These devices look like ordinary office telephones but they connect to a LAN network rather than the PABX system.

An advantage of the VoIP handset over the PC option is that it can provide access to the VoIP system regardless of the state of the user's PC. The PC does not have to be switched on at all times to receive calls. The problem with handsets is their lack of flexibility and upgradability to new standards. The handset software is proprietary so the user is locked into using the codecs that the vendor decides to support. However, if a PC was used as the VoIP platform then all that would be required is a software upgrade to support the new protocol. As time progresses though, handsets will be produced that will allow some re-programmability, due to advances in electronics and the need for companies to allow their hardware to remain valid in the future.

Another application of VoIP is for the telecommuter. Telecommuters require a data connection to their place of work so they can collaborate. At the same time they require voice connections to work not only so they can confer with colleagues but also to respond to external customer calls. By giving the telecommuter a VoIP telephone, the physical location of the worker can be transparent. This means that workers will only ever have to have one enterprise based telephone number, with the intelligence in getting to the end user located inside the VoIP system. This is effectively an enterprise equivalent to the 'virtual second line' technology.

3.3 Economic Forces (Toll Bypass)

Toll Bypass is currently the main driving force behind the take up of VoIP. It takes advantage of an *arbitrage* situation with telecommunications pricing. Currently, data capacity is priced at a flat rate no matter the physical destination. The rate per bit is also much lower than the equivalent PSTN charge. Money can be therefore saved by calling distant (especially overseas) locations using data links rather than the traditional PSTN.

- In the enterprise, companies are able to Toll Bypass by using their existing data connections to transport voice calls rather than the PSTN network.
- VoIP also offers the opportunity for second tier carriers to offer cheap long distance phone calls to users. The carriers use the economic *arbitrage* situation to compete with traditional carriers.

Such use of VoIP does not necessarily effect the end user who may not be aware that the call is being carried over the data links. In this case VoIP is implemented only between the telephone switches and the move to VoIP technologies can be transparent to end-users.
This price *arbitrage* is a short term effect however, it will fade as market forces cause pricing structures to realign. Therefore, these particular economic forces can be considered a short term driver for VoIP.

4 Technologies

4.1 Overview

This section will cover the technologies that are behind the VoIP concept. VoIP requires a real-time stream of information to be encoded and sent over a packet network, so there are many complex issues that need to be addressed. Users of voice networks also expect extra features such as call holding and call forwarding, so these services must be supported by the VoIP protocols.

Standards will allow a common way for systems from many vendors to communicate, as well as allowing users to connect to anyone from anywhere. As always, there is not just one single standard. The main players that are offering protocols and standards are the International Telecommunications Union (ITU) with H.323 [3] and the Internet Engineering Task Force (IETF) with Session Initiation Protocol (SIP) [4,5,6,7]. The ITU H.323 is by far the most widely implemented protocol at this time. It has been adopted by most of the big players in telephony, but complexity issues make it more difficult for smaller vendors to use and compete with the major players. Hence other protocols, such as SIP, are being developed which are somewhat simpler and ultimately making it easier and cheaper to implement. Adoption of SIP in the market place over the last few months is accelerating as big companies realise the benefits of the protocol.

4.2 Network Elements

A VoIP system will generally comprise a number of functional elements. (Within this sub section H.323 based terms will be used). Users interface with a terminal. Controlling the connections between terminals is a device called a gatekeeper. Connection from the VoIP system to traditional phone systems occurs via a gateway. A conference between multiple calls may be provided via a conference bridge (otherwise known as a multipoint control unit - MCU).

4.2.1 Terminal

The terminal can be any end-point device that connects to a LAN and translates voice, video or data into a format suitable for transport over the network. A PC and associated software, or dedicated hardware/software combination, contained in a special handset could perform this function. Endpoint is a general term that describes a device that terminates a call. It usually interfaces to a human at either end of a call, but could also be as an example, a Voice Mail unit.

4.2.2 Gatekeeper

Gatekeepers perform many functions, in particular the gatekeeper is the mechanism for:

- Call control and call routing.
- Basic telephony services such as directory services.
- PBX functions (e.g. call transfer, call forwarding).
- Controlling VoIP bandwidth usage to assist with QoS and protect other critical network applications from VoIP traffic.
- Injection of overall system administration and security policies.

In real world implementations, the gatekeepers provide:

- Internet Service Providers having the ability to do billing for guaranteed bandwidth management and special service packages.
- Intranet managers having seamless interoperability between PBX dial plans and IP-based terminals.
- Network managers having rapid, easy-to-use interfaces to modify or update zone configuration when an individual on the network needs additional services.
- Multimedia call centres for customer service being able perform needs-based call routing and a variety of other automatic call distribution features.

4.2.3 Gateway

A Gateway provides translation of protocols for call setup and release, conversion of media formats and transfer of information between H.323/SIP and other networks. Thus a gateway can be used to connect VoIP with an analogue PSTN system. Gateways are optional if connections to other networks are not required as terminals can communicate with each other if they are on the same IP network.

4.2.4 Multi-point Control Unit

MCUs are used to handle conferencing between three or more end-points. This can be a stand-alone unit or integrated into a gateway, gatekeeper or terminal. There are two parts to its functionality, the Multi-point Controller (MC) which handles control and signalling for conference support and Multi-point Processor (MP) which receives data streams from end points and processes them for re-distribution to other end-points.

Multi-point conferences can be centralised or decentralised, using unicast and/or multicast methods of data distribution, see Figure 1. H.323 and SIP can support a mixture of these two conferencing modes.

- One advantage of centralised conferencing is that it may output multiple unicast connections and all data switching and conversions are handled by the MCU, making the terminals simpler and reducing the bandwidth demands on the system.
- The decentralised model would require the use of multicasting and smarter and more complicated terminals, with an increase in network traffic. In this case, the MCU would provide mostly only MC functionality but MP could also be provided for terminals that required it.

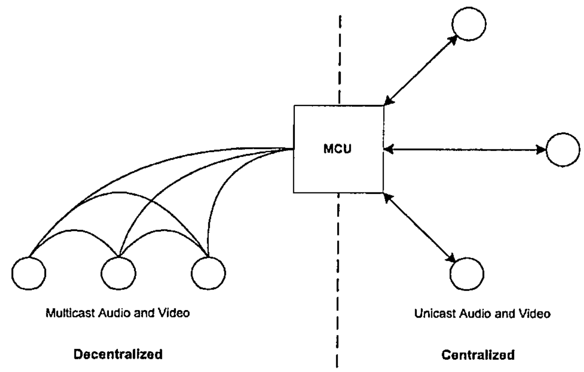

Figure 1 Conference Modes

4.3 ITU protocol H.323

H.323 is a binary protocol (ASN.1 notation/encoding) and broad in scope. It includes voice coding/decoding (codec) standards, stand-alone devices, VoIP embedded in personal computers (PCs), point to point and multi-point conferences and platform and application independence. In addition, H.323 addresses call control, multimedia management, bandwidth management and interfaces between LANs and other networks. H.323 is part of a larger H.320 communications standard. Figure 2 shows a typical H.323 Zone and includes elements such as the MCU, Gateway, Gatekeeper and Terminal.

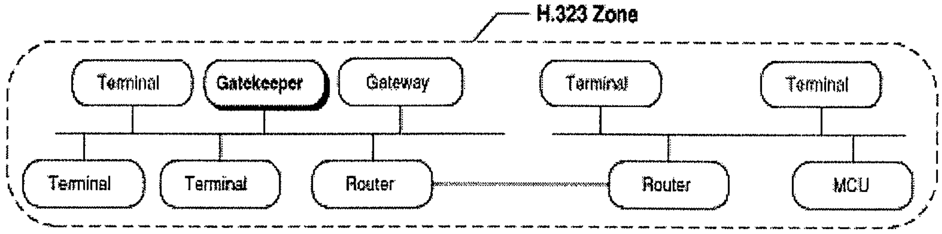

Figure 2 H.323 Devices

Version 1 of H.323 does not provide a guaranteed QoS. The current version of H.323 is Version 2 that supports connections to traditional switched networks.
The H.323 stack is shown in Figure 3. It shows the protocols and mechanisms for the connection and transport of VoIP data. H.225.0 Registration, Admission & Status (RAS) is the control signalling between an endpoint and a gatekeeper. H.225.0 (based on Q.931) handles call signalling between endpoints or endpoints and the gatekeeper/MCU. H.245 performs control signalling between endpoints or endpoint and gatekeeper/MCU to determine capabilities. Resource reservation protocol (RSVP) is used to prioritise and guarantee latency to specific IP traffic streams. Real-Time Protocol (RTP) is used for the transport of real time data such as audio and video over the network. Real time Transport Control Protocol (RTCP) provides information on the transmission and reception quality of data carried by RTP. These streams are sent unreliably on User Datagram Protocol (UDP), while the call signalling and control signals use reliable Transport Control Protocol (TCP).

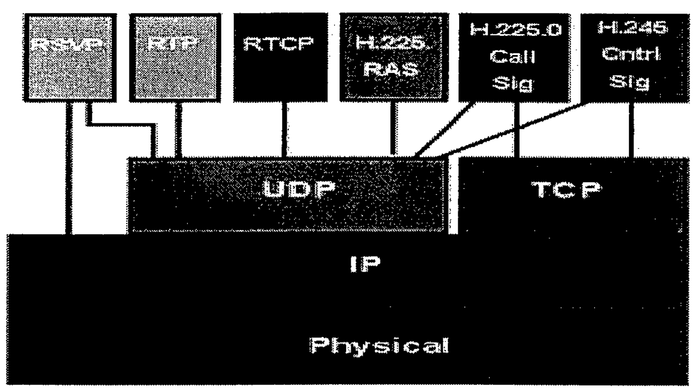

Figure 3 H.323 Protocol Stack

4.3.1 Terminal

The terminal is required to support a number of functions within the H.323 protocol as shown in Figure 4.

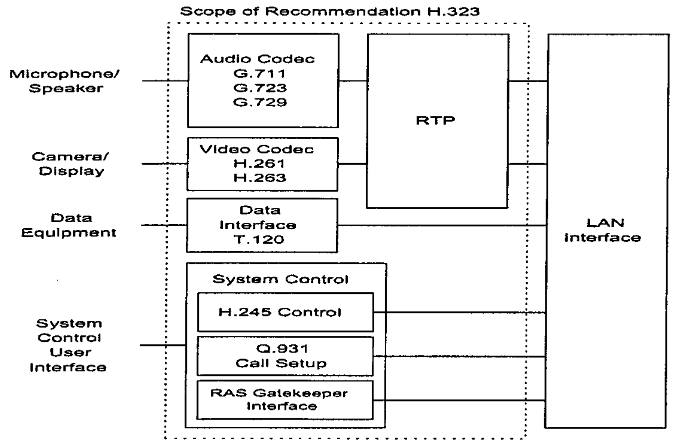

Figure 4 Terminal Standards

The terminal must support the audio format known as G.711, and can optionally support the following audio codecs: G.726, G.723.1, G.728, G.729 (see Table 5 on page 18). These audio codecs are ITU ratified codecs that are designed for various operating conditions. The choice of codec depends on the expected uses of the end terminals. Optional components are Video codecs: H.261 and H.263, T.120 data-conferencing protocols and MCU capabilities.

4.3.2 Gateway

A H.323 Gateway employs the family of standards shown in Figure 5. The shaded areas are parts defined by the standard.

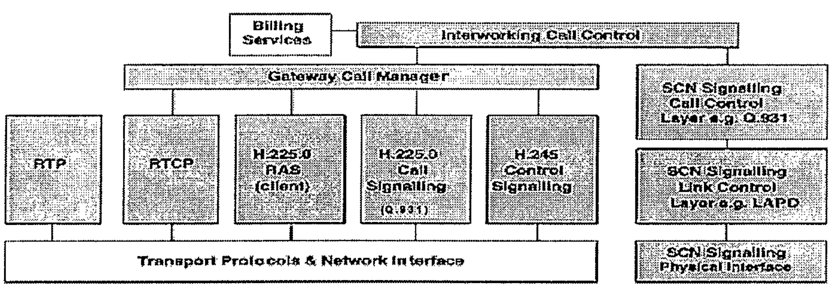

Figure 5 Gateway Standards

H.323 uses a variety of standards derived from two main sources, the IETF and ITU itself. The transport protocol (RTP) is defined by the IETF and it is used to provide a data stream over a packet network. The other standards are defined by the ITU and they provide the various standards relating to telephone functionality.

4.3.3 Gatekeeper

The gatekeeper is the most important component of a H.323 network. It is the central point for all calls within a zone and provides call control services between any terminals. While H.323 makes this optional, without it calls are peer-to-peer and can lead to an overloaded network and hence little or no QoS.

If a gatekeeper is in a system it is required to provide mandatory functions which are:

- Address translation (or routing).
- Admission control.
- Bandwidth control (via requests).
- Manage a Zone or Zones of H.323 devices.
- Handle backup and load balancing.

It is a software application, hence can be implemented on a PC platform, although it may also be integrated into a gateway or even a terminal unit. There are some optional functions that a gatekeeper can have such as billing/directory services, security, call policy/management and call control signalling, such as direct handling of Q.931 signalling between end points as seen in Figure 6. The gatekeeper modifies the control signalling, while the data streams are untouched.

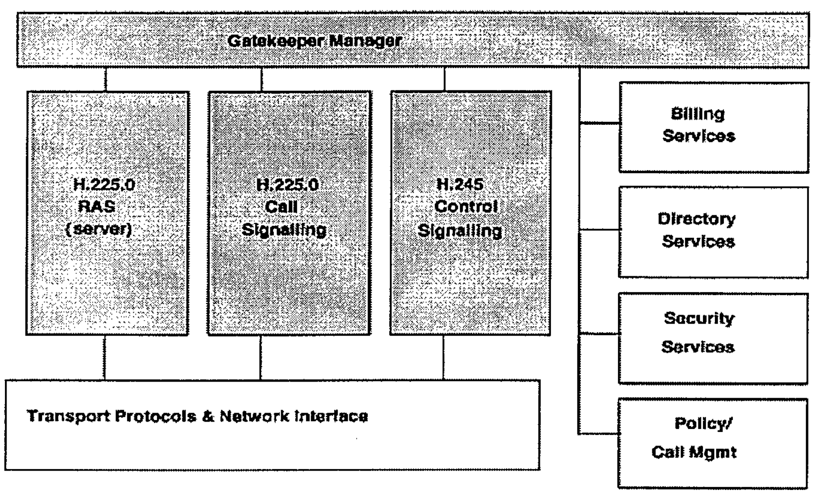

Figure 6 Gatekeeper Standards

4.3.4 H.323 Version 2

The addition of H.235 to the H.323 version 2 standard gives security and authentication features, such as, the use of passwords for registration with a gatekeeper. Other services such as call transfer and call forwarding are provided by H.450.x. Adding fast call set-up enables the bypassing of some set up messages. Lastly, the ability to specify alternative gatekeepers to endpoints adds more flexibility.

4.4 IETF protocol SIP

SIP is a text-based protocol similar to the Hyper Text Transfer Protocol (HTTP) used in the world wide web. It is not yet developed to the stage that H.323 is and its specification is lightweight compared to H.323. This is one of its strengths, enabling simpler processing modules to handle data processing leading to economic benefit, especially for smaller companies. Currently only prototype SIP terminals have been fielded.

The SIP protocol stack is shown in Figure 7 and the RSVP, RTP and RTCP protocols function in the same fashion as described in the H.323 Stack above.

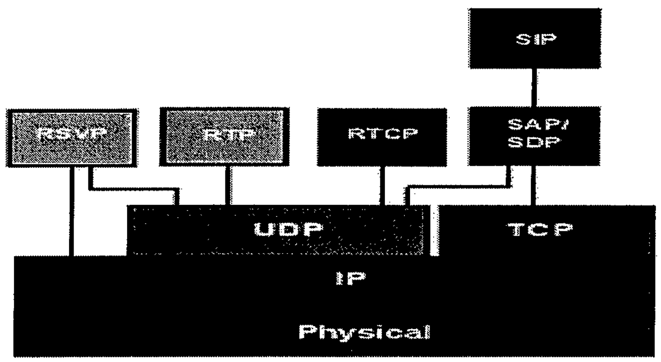

Figure 7 SIP Protocol Stack

SIP handles basic and supplementary services to create, modify and terminate multi-media sessions. Service Advertising Protocol (SAP) is required for advertising multicast conferences and other multicast sessions, while Session Description Protocol (SDP) describes multimedia sessions. SIP employs the SIP Server to set-up and control client connections, with similar functionality to H.323 gatekeeper functions.

SIP has been developed with modularity and extensibility in mind and can be seen via the call control functionality. SIP uses one protocol, while H.323 splits functionality across H.450, RAS, H.245 for IP plus Q.931 for PSTN connections. This is seen when comparing the SIP stack in Figure 7 and the H.323 stack in Figure 3. SIP allows for a diverse range of codecs, which can include H.323 codecs, that are registered with IANA (Internet Assigned Numbers Authority [8]), or even privately assigned ones. While SIP call set-up can use TCP, generally it uses only UDP which enables faster call set-up. (Originally H.323 used TCP for this, but by version 3, it too will use UDP as this provides a faster method [9].)

4.4.1 SIP Services

There are three basic services, the user agent, the proxy and the redirect. The user agent is just the terminal/client, while the proxy server does the call set-ups and controls client connections with similar functionality to a H.323 gatekeeper function. The redirect server is different in that it merely passes call control to another server and does not hold call parameters (as does the proxy) once passed on. There is no formal gateway as in H.323 and it is expected the SIP terminal will handle the actions required.

4.4.2 Multipoint Conferencing in SIP

SIP does not have or use an MCU unit (as in H.323) to co-ordinate conference calls. In SIP a bridge is formed between callers. As with the MCU, SIP allows conferencing by data being sent between all users (full-mesh) or one user can be the master and mix all other audio signals and re-send (bridged). Finally, to reduce bandwidth, multicast transmissions can be used.

4.4.3 SIP's main message constructs

- **Invite** invites a user to a conference
- **Bye** terminates a connection between two users
- **Cancel** request cancels a pending request with the same Call-ID
- **Options** signals information about capabilities
- **Status** informs the server about the progress of signaling
- **Ack** is used as a response in reliable message exchanging
- **Register** conveys location information to a SIP server

The latest IETF draft Request for Comment for SIP can be viewed at [10].

4.5 Comparison of H.323 and SIP

Technical and Business comparisons are shown in Table 2 and Table 3 respectively (from [6]). A more detailed technical comparison, including version 3 of H.323 is shown in Table 4 (from [11]). As can be seen from the tables, H.323 is more established and offers many more options for VoIP, while SIP is an emerging standard that is better suited to simpler devices, such as a Personal Digital Assistant. One strong point for H.323 is that it is totally backward compatible [11], while SIP cannot guarantee that some earlier functions may not be replaced with better and more efficient means to perform that functionality.

Some examples of each protocols call setup stage is shown in Appendix A. The activity in the commercial world is quite high and consequently SIP, which 6 months ago seemed like only a possible contender with H.323, is now emerging as a true alternative. To this end, new groups have been set-up to handle H.323 and SIP interoperability[12], but for now connectivity can be handled to some extent via a translating gateway. For further SIP related material, view [13] while for H.323, information must be purchased from the ITU in [14]. For more comparisons look at references given in [15].

Table 2 Technical Comparison

H.323	SIP
Complex	Simple
Hundreds of elements	42 headers
Binary ASN.1 coding	Text Based
Intertwined protocols	Modular protocols
Many options	Simple operations

Table 3 Business Comparison [6]

H.323	SIP
Established - deployed in many areas and equipment manufactures, so immediate interoperability will not be an issue	New Protocol - units just coming to market now
Can offer reliable solutions and avoid single point failures	Can offer reliable solutions and avoid single point failures
Can offer tiered services to the customer	Easily deployable - hence can generate revenues quickly
Has industry support from Microsoft, AOL, even browsers support it	Backed by Cisco, 3Com, Ericsson, Siemens, Motorla and others.
Differentiated services support, such as bit rate and delay negotiation	Not supported
Suitable for Devices with large processing capacity	Suitable for lightweight devices, (e.g., PDAs, mobiles), as well as desktops
Good business opportunities since it suitable for many points in a network.	Good business opportunities with a higher volume.

Table 4 More detailed Technical comparison

	H323.v1	H323.v2	H323.v3	SIP
FUNCTIONALITY				
CALL CONTROL SERVICES:				
Call Holding	No	Yes	Yes	Yes
Call Transfer	No	Yes	Yes	Yes
Call Forwarding	No	Yes	Yes	Yes
Call Waiting	No	Yes	Yes	Yes
ADVANCED FEATURES:				
Third Party Control	No	No	No	Yes
Conference	Yes	Yes	Yes	Yes
Click-for-Dial	Yes	Yes	Yes	Yes
Capability Exchange	Yes&Better	Yes&Better	Yes&Better	Yes
QUALITY OF SERVICE:				
Call Setup Delay	6-7 RT*	3-4 RT*	2-3 RT*	2-3 RT*
RELIABILITY:				
Packet Loss Recovery	Through TCP	Through TCP	Better	Better
Fault Detection	Yes	Yes	Yes	Yes
Fault Tolerance	N/A	N/A	Better	Good
MANAGEABILTY				
Admission Control	Yes	Yes	Yes	No
Policy Control	Yes	Yes	Yes	No
Resource Reservation	No	No	No	No
SCALABILTY				
Complexity	More	More	More	Less
Server Processing	Stateful	Stateful	Stateful or Stateless	Stateful or Stateless
Inter-Server Communication	No	No	Yes	Yes
FLEXIBILITY				
Transport Protocol Neutrality	TCP	TCP	TCP/UDP	TCP/UDP
Extensibility of Functionality		Vendor Specified		Yes, IANA
Ease of Customisation	Harder	Harder	Harder	Easier
INTEROPERABILTY				
Version Compatibility	N/A	Yes	Yes	Unknown
SCN Signaling Interoperability	Better	Better	Better	Worse
EASE OF IMPLEMENTATION				
Protocol Encoding	Binary	Binary	Binary	Text

* RT is Round Trip

4.6 Gateway Control Protocols

What has been described so far is VoIP in an IP network environment, whilst we have shown that for example, a gateway can link VoIP to an SCN, e.g. PSTN [6]. Calling or interfacing can be performed by the Media Gateway Controller (MGC), which provides the architecture for call signalling/control between the two systems, that is, via the Gatekeeper in H.323 and SIP Proxy in SIP. Again two methods were proposed, MGCP (Media Gateway Control Protocol) and MEGACO, and the evolution of these standards is shown in Figure 8.

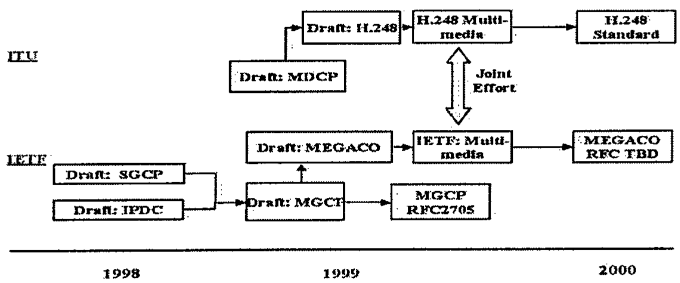

Figure 8 Evolution in Gateway Control Protocols

MEGACO is becoming H.248 via a joint effort between ITU and IETF, whilst MGCP is already defined by RFCs and hence has gained broader market support. Technically MGCP is simple with a straightforward command set, whilst MEGACO is complex but has a more flexible command set. Figure 9 shows the use of these protocols for H.323 and SIP.

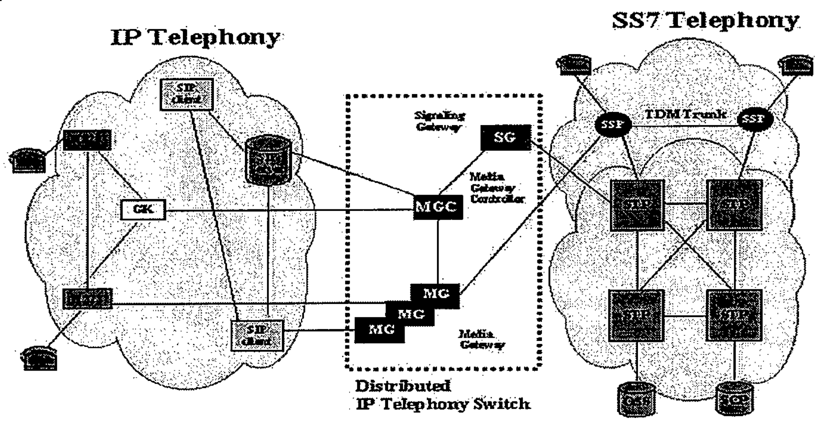

Figure 9 Gateway Approach

4.7 Fax or FoIP (Fax over IP)

Both the ITU and IETF are working on two standards, T.37 (for store and forward operation) and T.38 (for real time fax connections). For H.323, T.38 has been selected as the standard to use. Fax in its basic form is digital, but is converted to analogue for connection via PSTN. New fax machines would use the digital data and packetize it, but for legacy units, a packet Interworking Functional (IWF) unit (or gateway) can be employed to convert analogue to packet and vice versa, as shown in Figure 10. QoS issues stem from the same problems as for VoIP, as described in 3.2. Delay through the network causes data skewing on fax units, since faxes must work synchronously (or close to it) for intelligible output [16].

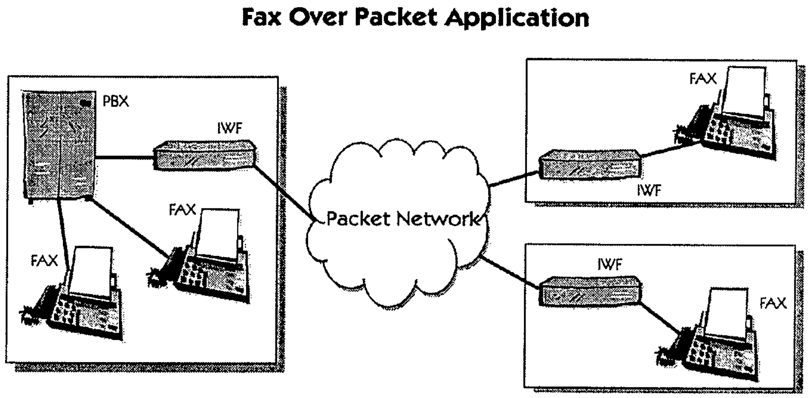

Figure 10 Fax over IP

4.8 Speakeasy

The interconnection of Speakeasy voice encryption units over VoIP presents some problems. Speakeasy encryption is digital using either a high rate voice codec for ISDN lines or a low rate codec for normal, analogue, PSTN connections. For analogue connections the digital stream is converted to analogue form via Fax type modems. A key concern would be the situation where the carrier (or the Defence Voice Network in its pseudo-carrier role) sought to employ VoIP within the core of the network without the participation of the Speakeasy end user.

Current Defence Voice Network identifies Speakeasy (and other data modems) as not suitable for compression. It is interesting to consider the impact of the number of Speakeasy/data modem calls on the business case for VoIP. A short informal examination of calls showed a high proportion (38 out of 243) of such calls being conducted on the Defence Voice Network [17].

In the case of ISDN, once encrypted, the 64 kbps digital stream no longer represents a voice call and cannot be sensibly processed by the low rate voice codecs for passage over VoIP. This issue is as relevant to Speakeasy/VoIP as it is to any in-network processing (compression/speech activity detection etc). Speakeasy negotiates, during signalling, a data channel (no compression permitted) rather than a voice channel which might be subject to network initiated compression (or VoIP).

In the case of PSTN mode there may be some scope for the Speakeasy connection to be recognised as a Fax connection and the real time fax over IP standards invoked. This would need to be investigated, but there is a concern that while Speakeasy may have the same modem waveform, the connection would not be recognised as a Fax connection because of a lack of Fax handshaking. End to end latency impacting on cryptographic synchronisation may be an issue.

The manufacturers of Speakeasy has already developed a device which can demodulate the Speakeasy modem waveform and provide an encrypted digital stream.

It would be feasible to use this device at the Speakeasy itself to provide an encrypted stream for packaging into an RTP IP stream for passage over an IP network. Whether one could call this VoIP is arguable.

5 End User Aspects

5.1 Quality Issues

Users of the telephone have come to expect a high voice quality with little delay as standard. This expectation is an issue for VoIP applications as only by using higher bandwidth codecs and QoS networks can this level of service be provided. This leads to a trade off for carriers, or possibly for the end user. They will have to make a value judgement about the cost versus quality of their voice transmissions. You can see this effect currently happening with the Internet Phone [18] market. People can use the Internet phone services for a substantial reduction in price as compared to normal carriers, but the voice quality of their conversations drops due to the unreliable nature of the Internet and the low bandwidth codecs used. In the future customers will be able to select the type of service they want, based on how they value the voice quality for that phone call.

A key determinant in voice quality is codec choice (see Table 5), but network performance will have a substantial impact on quality. These network factors are latency, jitter (latency variation), packet loss and echo compensation.

- Latency is basically the delay between end users. A delay of 100msec or less is considered desirable, whereas 200msec or greater is noticeable and will cause people to switch to half-duplex conversation. Delay is made up of three elements, accumulation delay or algorithmic delay, processing (packetisation) delay and network delay. Accumulation delay is caused because a "frame" comprising many voice samples must be collected before processing can be carried out on the frame. Processing delay is caused by the processing (compressing) of a frame and collection of encoded samples into a packet for transmission. Often multiple small packets are collected in a single larger packet to reduce network overhead (the ratio of headers to useful data). Lastly, network delay is the time taken for the packet to be passed across the network to the recipient. This is partly determined by the physical distance to be traversed and the capacity (bit rate) of links along the way, but also due to packets being queued awaiting their turn on each link.
- Jitter is introduced when packets traverse different paths on the network or because packets suffer different queuing delays in the network, because of variations in competing traffic. Buffers at the receiving terminal are employed to remove jitter, but this leads to greater latencies on the channel. Adaptive means are often employed to vary buffer size dependant upon the amount of jitter in an effort to minimise the impact on latency.
- Since packets are sent using UDP, an unreliable protocol, codecs must be able to handle some packet loss, e.g. by interpolation, but a loss of 5% or more is usually noticeable. The amount of packet loss a codec can handle before voice performance

is degraded determines the robustness of the protocol. Speech is continuous, but packets may arrive out of order so protocols must be in place to prevent sequence errors. If a packet has not arrived at a receiving terminal before it is due to be played out to the listener, it is effectively a lost packet.

- Echo is the returning of the speaker's voice back into the speaker's headset from the far end of the connection. Echo is tolerable if a round trip delay is 50msec or less but above this becomes disturbing. Echo cancellation techniques are required to remove this re-bounding, usually by implementing some form of a digital filter on the receive path from the packet network (essentially, a delayed version of the speakers input is subtracted from what is received).

Because of the impact of network performance on voice quality, there is a need for support to guarantee QoS from the network. For a more involved discussion on QoS and ways to implement it, please refer to the "New protocols for switching and traffic control in IP networks" paper from this series.

Table 5 Codec Choice

Codec	Audio bit rate	Complexity	Quality	Digitising Delay
G.711 PCM	48/56/64	N/A	Very Good	Negligible
G.726 ADPCM	40/32/24	Low(8 MIPS)	Good (40K) to Poor (16K)	0.125uS
G.722 Sub-band ADPCM	48/56/64	As above	Good	<1mS
G.729 CS-ACELP	8	High (30 MIPS)	Good	10mS
G.729A CA-ACELP	8	Moderate	Fair	Low
G.723 MP-MLQ	6.4/5.3	Moderate-High (20 MIPS)	Good (6.4K) to Fair (5.3K)	High
G.723.1 MP-MLQ	6.4/5.3	As above	As above	30mS
G.728 LD-CELP	16	Very High (40 MIPS)	Good	2.5mS

5.2 Quality Measures

The ITU has proposed a measure of voice quality, including for VoIP networks (ITU-T G.107 [19] and discussed in [20]). This uses multiple factors, including those discussed earlier, to assess performance and allocate a rating R. The method of measurement of the factors are such that the measurements are additive in respect of the quality rating:

$$R = R_0 - I_s - I_d - I_e + A \quad \text{(from [20])}$$

Where

R: the perceived quality of the call

R_0: This is the fundamental rating of a system as affected only by noise (background and, for analogue services, within the circuit)

I_s: Impairments that occur simultaneously with the production of the voice signal (such as quantisation of the voice signal by the codec)

I_d: Delayed impairments caused by echo or by a loss in interactivity - fundamentally determined by the network

I_e: Impairments due to special equipment (such as the effects of packet loss)

and A: The access advantage of the system - a measure of the degree to which users are prepared to tolerate a reduction in quality as a price to pay for more convenient access than the traditional wire network.

The implication of this formula is that impairments can be traded off with no perceived change in the quality rating of a call. For instance it is possible to have a call with high network delay (high I_d) but compensated by low packet loss (I_e) rating equally with one with high packet loss but low delay. While the two calls would sound completely different to the ear, listeners would still rate them at the same quality.

6 Market Trends

Currently most telecommunications carriers are implementing VoIP or telephony over Asynchronous Transfer Mode [21] (ATM) as a push towards a packet based world. Examples are Telstra's "Data Mode of Operation" and Optus' "Integrated Convergent Optus Network". Both of these programs aim to change the carriers internal network so that they are entirely packet based. The telephony codecs being used on these internal networks are likely to be G.711, the PCM audio codec that operates at 56 or 64 kbps. This is mainly because it maintains the current circuit switched voice quality that consumers have come to expect. The beauty of carriers implementing packet networks is that they can transparently change the codecs they use, so they can leverage the codec that offers them the best result.

Carriers are also likely to implement ATM as their internal network infrastructure. This is for two reasons, the biggest being that ATM was developed for and by carriers so it is most suited towards their needs. From the perspective of this paper, such an implementation is then well placed to provide the network QoS that is vital in the proper implementation of VoIP systems.

In the consumer market (end users and enterprises), the current trend is still towards H.323 as the VoIP standard. This is mainly due to the large installed base of H.323 clients. SIP is a simpler protocol and seems to be gaining support due to its better feature set, but only time will tell if it will be widely adopted. The current audio codec standard in the H.323 market is G.723.1 that is a speech codec for 5.3 and 6.4 kbps voice streams. This codec provides good voice quality at the lowest data rate.

As well as offering basic voice services via VoIP, consumers will also be able to access new features such as caller-id and conference calls between an arbitrary number of people. More importantly, all of these features will be much easier to use because of the CTI aspect of VoIP. In addition, features that do not make sense in the current voice networks such as web based call centres will be created as the technologies are invented.

7 References

[1] Virtual second line, http://www.telstra.com.au/research/h_virtua.htm, March 2000.
[2] Voice Over Packet Networks: An Enterprising Overview, http://www.gartner.com/webletter/cisco/article3/article3.html, March 2000.
[3] ITU-T Rec. H.323, "Visual Telephone Systems and Terminal Equipment for Local Area Networks which provide a non-guaranteed Quality of Service", 1996.
[4] A primer on H.323 series standard, http://www.databeam.com/h323/h323primer.html, March 2000.
[5] ITU-T Recommendation H.323, Feb 1998 CD.
[6] Evolving VoIP Protocol Standards, H.323 vrs SIP- Prasad Kallar of Trillium and MGCP vrs MEGACO- Philip Fote of Radisys, Netseminar of 14/3/2000 sponsored by Intel. http://www.pulver.com
[7] Voice Over Packet, White paper, Edward B. Morgan of Telogy Networks Inc. http://www.telogy.com/
[8] http://www.iana.org/numbers.htm Lists Audio codecs
[9] http://www.cs.columbia.edu/~hgs/papers/Eyer0004_Predicting.pdf
[10] http://www.cs.columbia.edu/~hgs/sip/drafts/draft-ietf-sip-2543bis-00.pdf
[11] Comparison of H.323 and SIP for IP Telephony Signaling, by Ismail Dalgic, Hanlin Fang http://www.cs.columbia.edu/~hgs/sip/papers.html
[12] IMTC (http://www.imtc.org/) and ETSI (http://www.etsi.org/tiphon/)
[13] http://www.cs.columbia.edu/~hgs/sip and http://www.ietf.org
[14] http://www.itu.int/publications/itu-t/itutrec.htm
[15] http://keskus.hut.fi/tutkimus/ipana/paperit/sip.pdf and http://www.hut.fi/~lhuovine/study/iwork99/voip.html
[16] Fax Over Packet, White Paper, Edward B. Morgan of Telogy Networks Inc. http://www.telogy.com/
[17] Informal examination on afternoon of 23 June 2000 by Mr Bernie Cass, DISC.
[18] Net2Phone , www.net2phone.com/
[19] "the E-Model, a Computational Model for Use in Transmission Planning", ITU-T Recommendation G.107, December 1998.
[20] Quality bounds for packetized voice transport by D. De Vleeschauwer, J. Jassen, G. H. Petit and F. Poppe, Alcatel Telecommunications Review, 2000
[21] The Multiservice IP Carrier Network, http://www.netreference.com/PublishedArchive/WhitePapers/IPCarrier_wp/IPCarrier_wp.html#3, March 2000.

Other relevant reading:
[1] Beyond Dial Tone: Opportunities for Value in IP Telephony, http://www.brooktrout.com/whitepaper/iptel_value.htm, March 2000.

[2] RADVision: H.323 Center: Handbook Library: Gatekeeper, http://www.radvision.com/info/handbook/gatekeep/, March 2000.
[3] IP Telephony- White Paper, Signaling by Bjarne Munch of Ericsson Australia http://www.ericsson.com/datacom/emedia/ip_telephony.pdf

Appendix 1 - Examples of Call set-up for H.323 and SIP

Typical H.323 call

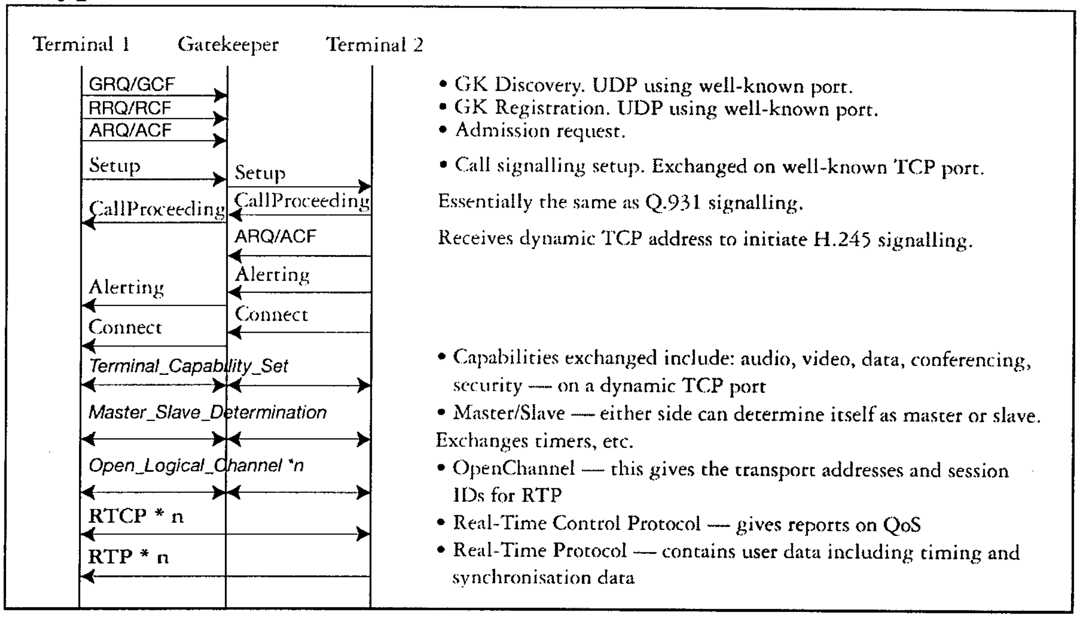

Figure 11

This shows H.323 in its usual mode communicating with the use of a gatekeeper. The audio Media Stream uses RTP*n where n is a port number.

Peer to Peer call

In the scenario shown in Figure 12 neither endpoint is registered to a Gatekeeper. The two endpoints communicate directly. Endpoint 1 (calling endpoint) sends the Setup (1) message to the well-known Call Signalling Channel TSAP Identifier of Endpoint 2. Endpoint 2 responds with the Connect (4) message which contains an H.245 Control Channel Transport Address for use in H.245 signalling.

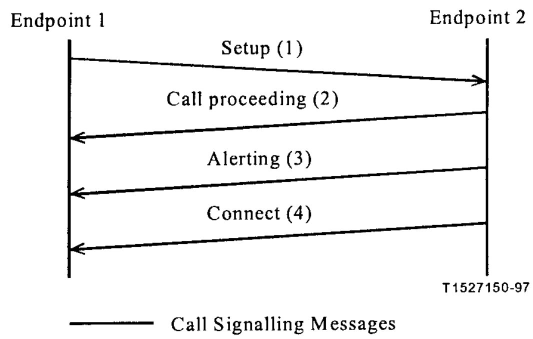

Figure 12

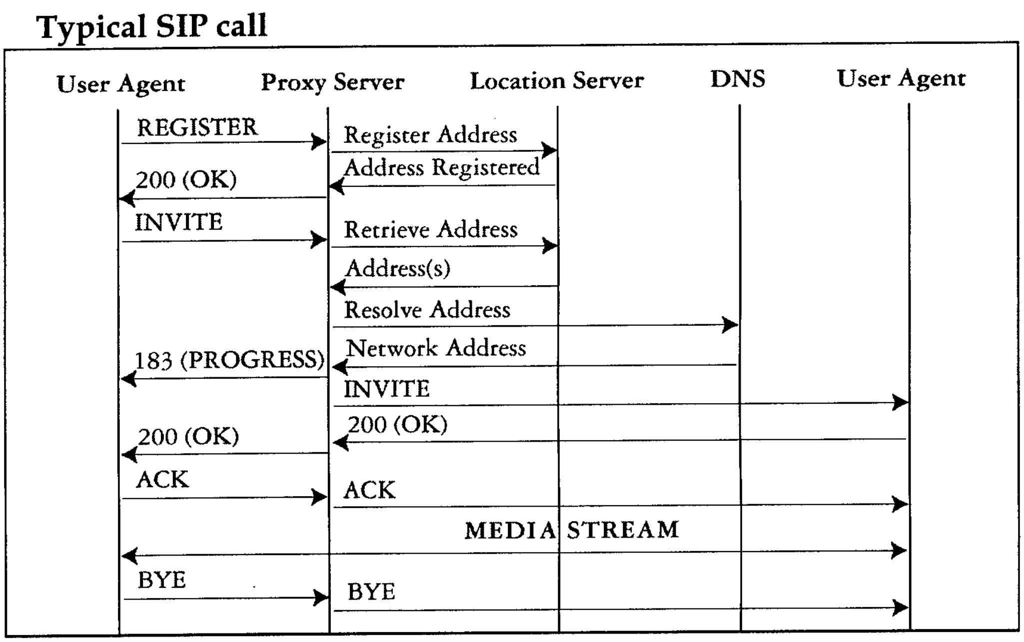

Figure 13

This shows SIP's ease in using a proxy (giving extra security) and address resolution via DNS. For audio, the media stream is RTP as it is for H.323.

DISTRIBUTION LIST

IP Convergence in Global Telecommunications

-

Voice over Internet Protocol (VoIP)

Ian Zahorujko, Alfred Reynolds and Bill Blair

AUSTRALIA

DEFENCE ORGANISATION

Task Sponsor

Director General C3I Development
PD JP 2047 (CWAN)
PD JP 2068 (NOC)
PD JP 2061 (EXC3ITE)

S&T Program

Chief Defence Scientist
FAS Science Policy
AS Science Corporate Management
Director General Science Policy Development
} shared copy
Counsellor Defence Science, London (Doc Data Sheet)
Counsellor Defence Science, Washington (Doc Data Sheet)
Scientific Adviser to MRDC Thailand (Doc Data Sheet)
Scientific Adviser Policy and Command
Navy Scientific Adviser (Doc Data Sheet and distribution list only)

Scientific Adviser - Army (Doc Data Sheet and distribution list only)

Air Force Scientific Adviser
Director Trials

Aeronautical and Maritime Research Laboratory
Director

Electronics and Surveillance Research Laboratory
Director (Doc Data Sheet and distribution list only)

Chief of Communications Division
Research Leader Military Information Networks
Head Network Requirements
Head Network Architecture
Head Wireless Systems
Research Leader Secure Communications
Head Intelligent Networks
Research Leader Military Information Enterprise Branch
Alfred Reynolds
Ian Zahurujko
Bill Blair

DSTO Library and Archives

Library Fishermans Bend (Doc Data Sheet)
Library Maribyrnong (Doc Data Sheet)
Library Salisbury
Australian Archives
Library, MOD, Pyrmont (Doc Data sheet only)

US Defense Technical Information Center, 2 copies
UK Defence Research Information Centre, 2 copies
Canada Defence Scientific Information Service, 1 copy
NZ Defence Information Centre, 1 copy
National Library of Australia, 1 copy

Capability Development Division

Director General Maritime Development (Doc Data Sheet only)
Director General Aerospace Development (Doc Data Sheet only)

Knowledge Staff

Director General Command, Control, Communications and Computers (DGC4) (Doc Data Sheet only)
Director General Intelligence, Surveillance, Reconnaissance, and Electronic Warfare (DGISREW)R1-3-A142 CANBERRA ACT 2600 (Doc Data Sheet only)
Director General Defence Knowledge Improvement Team (DGDKNIT) R1-5-A165, CANBERRA ACT 2600 (Doc Data Sheet only)
Director Communications and Computing Networks (full copy)

Army

ABCA Standardisation Officer, Puckapunyal, (4 copies)
SO (Science), DJFHQ(L), MILPO Enoggera, Queensland 4051
NAPOC QWG Engineer NBCD c/- DENGRS-A, HQ Engineer Centre Liverpool Military Area, NSW 2174 (Doc Data Sheet only)

Intelligence Program

DGSTA Defence Intelligence Organisation
Manager, Information Centre, Defence Intelligence Organisation

Defence Information Systems

DGIS-DIS
DGCIPP-DIS
D-DISC

Corporate Support Program

OIC TRS, Defence Regional Library, Canberra

UNIVERSITIES AND COLLEGES

Australian Defence Force Academy
Library
Head of Aerospace and Mechanical Engineering
Serials Section (M list), Deakin University Library, Geelong, 3217
Hargrave Library, Monash University (Doc Data Sheet only)

Librarian, Flinders University

OTHER ORGANISATIONS

NASA (Canberra)
AusInfo
State Library of South Australia
Parliamentary Library, South Australia

OUTSIDE AUSTRALIA

ABSTRACTING AND INFORMATION ORGANISATIONS

Library, Chemical Abstracts Reference Service
Engineering Societies Library, US
Materials Information, Cambridge Scientific Abstracts, US
Documents Librarian, The Center for Research Libraries, US

INFORMATION EXCHANGE AGREEMENT PARTNERS

Acquisitions Unit, Science Reference and Information Service, UK
Library - Exchange Desk, National Institute of Standards and Technology, US

SPARES (5 copies)

Total number of copies: 62

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA	1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

2. TITLE	3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)
IP Convergence in Global Telecommunications - Voice over Internet Protocol (VoIP)	Document (U) Title (U) Abstract (U)
4. AUTHOR(S)	5. CORPORATE AUTHOR
Ian Zahorujko, Alfred Reynolds and Bill Blair	Electronics and Surveillance Research Laboratory PO Box 1500 Salisbury SA 5108 Australia

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-TR-1039	AR-011-569	Technical Report	September 2000

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
E8709/011/0015	JNT99/141	DG C3ID	24	24

13. URL ON THE WORLDWIDE WEB	14. RELEASE AUTHORITY
http://www.dsto.defence.gov.au/corporate/reports/DSTO-TR-1039.pdf	Chief, Communications Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE CENTRE, DIS NETWORK OFFICE, DEPT OF DEFENCE, CAMPBELL PARK OFFICES, CANBERRA ACT 2600

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CASUAL ANNOUNCEMENT Yes

18. DEFTEST DESCRIPTORS

Communication Networks, IP Networks, Network Architecture

19. ABSTRACT

A key application on any converged global network will be voice telephony. On an Internet Protocol (IP) network, this is given the generic title Voice over IP (VoIP). This paper examines the motivation behind VoIP and the standards being deployed in support of the application. It discusses the factors that determine the voice quality to users, and measures that can be made. The impact of VoIP on Speakeasy is given some limited consideration.

TECHNICAL REPORT DSTO-TR-1039 AR-011-569 SEPTEMBER 2000

DEPARTMENT OF DEFENCE
DEFENCE SCIENCE & TECHNOLOGY ORGANISATION
DSTO

ELECTRONICS AND SURVEILLANCE RESEARCH LABORATORY
PO BOX 1500 SALISBURY SOUTH AUSTRALIA, 5108
AUSTRALIA, TELEPHONE (08) 8259 5555